

UNITED STATES PATENT APPLICATION FOR

MOTION DETECTION USING MULTI-RESOLUTION IMAGE  
PROCESSING

INVENTORS:

Tong Sen LIEW

Tatt Wei HO

## MOTION DETECTION USING MULTI-RESOLUTION IMAGE PROCESSING

### TECHNICAL FIELD

Embodiments in accordance with the present invention relate to the  
5 field of image processing. Specifically, embodiments in accordance with  
the present invention relate to detecting motion involving decomposing  
images into multiple frequency bands.

### BACKGROUND

10 Recently, electronic devices have been developed that are able to  
track the movement of the electronic device by comparing images  
sampled by the electronic device at different points in time. For example,  
optical computer mice compare images of a surface that the mouse is  
moved over to determine motion of the optical mouse.

15

In one conventional system, the electronic device has a small  
light-emitting diode (LED) that emits light that reflects light off almost  
any surface and back to a complimentary metal oxide semiconductor  
(CMOS) sensor in the electronic device. This produces in a two-  
20 dimensional image of a portion of the surface. The CMOS sensor  
digitizes the two-dimensional image and sends each digitized image to a  
digital signal processor (DSP), which compares the current image with  
a previously captured reference image to determine the motion of the  
electronic device.

25

Such optical navigation has many advantages, among which are  
the lack of moving parts that accumulate dirt and suffer mechanical

wear and tear of use. Another advantage of an optical mouse is that it does not need a mouse pad, since it is generally capable of navigating upon arbitrary surfaces, so long as they are not optically featureless. Optical navigation works quite well for a variety of surfaces. However,  
5 due to the nature of some surfaces, the electronic device tracks well when moved on some directions across the surface, but has difficulty tracking when moved in another direction across the surface.

As optical navigation is based on a comparison of images,  
10 repetitive features in a surface can make it difficult to properly compare the images. For example, some optical mice encounter difficulties when operating on some wood grained surfaces. Grainy surfaces may also present problems, depending upon feature size, as well as orientation.

15 Thus, one problem with conventional methods to detect motion using image processing methods is the difficulty in working with a wide spectrum of images. Another problem with conventional methods to detect motion using image processing methods is the difficulty in detecting motion with respect to some surfaces.

## DISCLOSURE OF THE INVENTION

The present invention pertains to a device to detect motion using sub-band image processing. An embodiment in accordance with the invention is a device having decompose logic that decomposes an input  
5 image into composite images that comprise different frequency bands of the input image. The device also has storage coupled to the decompose logic to store composite images as reference images for comparison with a later input image. The device further has comparison logic to compare the composite images with the reference images to produce preliminary  
10 motion values for the different frequency bands. The device also has logic to determine a final motion value from the preliminary motion values.

## BRIEF DESCRIPTION OF THE DRAWINGS

The accompanying drawings, which are incorporated in and form a part of this specification, illustrate embodiments according to the invention and, together with the description, serve to explain the principles of the invention:

Figure 1 is a flowchart illustrating a process of motion detection by decomposing images into frequency bands, in accordance with an embodiment of the present invention.

Figure 2 is a block diagram illustrating a system for detecting motion by decomposing images into frequency bands, in accordance with an embodiment of the present invention.

Figure 3 is a block diagram illustrating a system for detecting motion by decomposing images into frequency bands with a filter bank, in accordance with an embodiment of the present invention.

Figure 4 is a flowchart illustrating a process of detecting motion by decomposing images into frequency bands using a Discrete Wavelet Transform, in accordance with an embodiment of the present invention.

Figure 5A is a diagram illustrating frequency bands decomposed from an image by an embodiment in accordance with the present invention.

Figure 5B is a diagram illustrating “x” and “y” component frequency bands decomposed from an image by an embodiment in accordance with the present invention.

## DETAILED DESCRIPTION OF THE INVENTION

In the following detailed description of embodiments of the present invention, numerous specific details are set forth in order to provide a thorough understanding of the present invention. However,

5   embodiments of the present invention may be practiced without these specific details or by using alternative elements or methods. In other instances well known methods, procedures, components, and circuits have not been described in detail as not to unnecessarily obscure aspects of the present invention.

10

Embodiments in accordance with the present invention detect motion of a device moved across a surface for which detecting differences between features is difficult for conventional techniques. For example, embodiments of the present invention are able to filter out  
15   repetitive information that could interfere with accurately determining motion of an optical mouse moved across a surface if the repetitive information were not removed.

Figure 1 is a flowchart illustrating a process 100 of detecting  
20   motion by breaking images into frequency bands. Steps of process 100 may be stored as instructions on a computer readable medium and executed on a general-purpose processor coupled thereto by a bus. In step 110, an input image is received. However, process 100 is not so limited. In one embodiment in accordance with the present invention,  
25   the input image is received from a CMOS sensor by a digital signal-processing component of an optical mouse. However, the present invention is not so limited.

In step 120, the input image is decomposed into multiple frequency band images. Each frequency band image contains a portion of the frequency information of the input image for a range of frequencies. In general, the frequency band images each focus on a separate range of frequencies. However, there may be some overlap in frequencies between two or more of the frequency band images. For example, the nature of the image decomposition may not transform the input image into the ideal frequency band. Moreover, the present invention is not precluded from deliberately having overlap in the frequency band of two or more frequency band images.

The minimum number of frequency band images is two. The maximum is not limited to any number. The number of frequency band images into which the input image is decomposed typically affects the results of the final motion detection. The results typically improve with a greater number of frequency band images, at least if only a few are used. Therefore, the number of frequency band images into which the input image is decomposed may be selected to achieve desired results.

20

In step 125, the frequency band images are stored for use as reference band images in a later iteration of process 100. For example, the reference band images are used when processing the next input image received in step 110. For each frequency band image produced in step 120, there is a corresponding reference frequency band image that contains information from the input image for the same frequency band.



In step 130, the frequency band images produced in step 120 are compared to reference frequency band images that were stored in step 125 of a previous iteration of process 100. For example, each frequency band image is compared to the corresponding reference frequency band image having the same range of frequency information. The comparison may be a cross-correlation, but the present invention is not so limited.

Although the images are separated by frequency, the images are processed in step 130 in the spatial domain (e.g., by cross-correlation), in some embodiments of the present invention. However, the invention is not be limited to processing in the spatial domain. In some embodiments in accordance with the invention, the processing in step 130 is performed by other cross-correlation or non-cross-correlation methods in the frequency domain. In other embodiments in accordance with the invention, the processing in step 130 is performed by cross-correlation or non-cross-correlation methods in the spatial-frequency domain.

The result of the comparison of step 130 is a set of preliminary motion values. In one embodiment, a motion value comprises a change in an "x" and a "y" coordinate. There may be one or more such motion values generated for each of the comparisons. However, the present invention is not limited to the motion values being changes in "x" and "y" coordinate values.

In step 140, a final motion value is determined from the preliminary motion values. For example, the preliminary motion values may be summed with selected motion values given more weight than

others, but the present invention is not so limited. By giving greater weight to certain frequencies, greater or less emphasis is placed on a certain frequency range. For example, it may be determined that a certain frequency range may contain information that is more or less  
5 useful in determining motion.

Moreover, by placing less emphasis on a certain frequency band, repetitive information may be removed. This allows tracking to be based on fewer, more prominent features. Further, the weighting is based on  
10 the direction of motion or based on surface properties such as fineness of the grain, in embodiments in accordance with the present invention. For example, a frequency band which works well or poorly for the x-direction may not work well or poorly for the y-direction. This may be due to a factor such as grain typically running in a certain direction on a desk or  
15 depend on how close the grains are spaced on the desk, although there may be other factors as well. Thus, the weighting for the changes in x-values is different from the weighting for the changes in y-values, in an embodiment in accordance with the present invention. The amount of weighting is predetermined in one embodiment; however, the present  
20 invention is not so limited. In another embodiment, the weighting is dynamically determined during optical navigation by statistical processing of the previous images. The process 100 then returns to step 110 to process another input image.

25 Figure 2 is a block diagram illustrating a system 200 for detecting motion by decomposing an image into frequency bands, in accordance with an embodiment of the present invention. The system 200 includes

an image sensor 210. In one embodiment in accordance with the present invention, the image sensor 210 is a CMOS sensor. However, the present invention is not so limited.

5           The image sensor 210 sends images to the image decompose logic 220. Throughout this application, the image decompose logic 220 may also be referred to as transform logic. The image decompose logic 220 decomposes the images into frequency band images. Each frequency band image contains a portion of the frequency information of the input  
10 image for a range of frequencies. The image decompose logic 220 sends the frequency band images to the reference image storage 225 for use when processing a later input image. Further, the image decompose logic 220 sends the frequency band images to the image comparison logic 230.

15           The image comparison logic 230 performs a comparison of the frequency band images from the image decompose logic 220 with the reference images from the reference image storage 225. The output of the image comparison logic 230 is a motion value for each comparison.  
20 This results in a motion value for each frequency band. It will be understood that a motion value may be expressed in terms of, for example, an x-coordinate and a y-coordinate.

          The motion value logic 240 inputs the motion values from the  
25 image comparison logic 230 and computes a final motion value therefrom. The final motion value is output from the motion value logic.

In one embodiment in accordance with the invention, the image decompose logic 220 is implemented with a filter bank. A filter bank embodiment will be discussed in more detail with respect to Figure 3. In another embodiment in accordance with the invention, the image decompose logic 220 is capable of performing a Discrete Wavelet Transform. One embodiment of the present invention is a redundant Discrete Wavelet Transform, in which down-sampling is not performed. However, embodiments of the present invention are well-suited to all implementations of the Discrete Wavelet Transform and its variants. A Discrete Wavelet Transform embodiment will be discussed in more detail with respect to Figure 4.

The comparison logic 230 is capable of performing a cross-correlation in one embodiment, but the present invention is not so limited. The present invention is not limited to processing in the spatial domain. In some embodiments in accordance with the invention, the comparison logic 230 is capable of performing other cross-correlation or non-cross-correlation methods in the frequency domain.

The various logic in Figure 2 may be implemented in an application specific integrated circuit, in one embodiment. In another embodiment, the image comparison logic 230, the image decompose logic 220, and the motion value logic 240 are implemented by a processor executing instructions stored on a computer readable medium.

25

Figure 3 is a flow diagram illustrating a system 300 for detecting motion by decomposing images into frequency bands, in accordance with

an embodiment of the present invention. In the embodiment of Figure 3, a filter bank 320 is used to decompose the input image; however, the present invention is not limited to this technique for decomposing the input image. The system 300 includes an image sensor 210, which  
5 digitizes and sends an input image to the filter bank 320. In one embodiment in accordance with the present invention, the image sensor 210 is a CMOS sensor. However, the present invention is not so limited.

The filter banks 320 are used decompose the input image into  
10 frequency band images in this embodiment. The filters may include a low-pass filter, a high-pass filter, and band-pass filters. The present invention is not limited to any particular combination of filters. Thus, it is not required that there be a low-pass filter, or a band-pass filter, or a high-pass filter. This embodiment of the present invention filters the  
15 input image into at least two frequency band images (not depicted in Figure 3) that contain different frequency information. There is no maximum number of frequency band images. Figure 3 depicts the filter banks 320 generating "n" frequency band images.

20 The filter bank 320 sends its output to cross-correlation logic 340a-n. Each of the "n" frequency band images is compared to its corresponding reference frequency band image 330a-n by performing a cross-correlation, in this embodiment. Thus, there are "n" cross-correlations 340a-n depicted in Figure 3. The cross correlation is performed on a  
25 subset of the pixels, in embodiments in accordance with the invention. Computation time is saved and the circuitry may be less complex by performing the cross-correlation on a subset of the pixels in the image.

The results of the cross-correlation are motion values for each comparison. The motion values describe the difference in physical location between the current input image and the reference image. Thus, the motion values describe how far an optical mouse that captured  
5 the images has been moved by a user between the time the reference image and the current image were captured, in an optical mouse embodiment.

A weighting or averaging function 350 produces a final motion  
10 value, based on the preliminary motion values. In this embodiment, the final motion value comprises a delta-x and a delta-y value. In one embodiment, the delta-x and delta-y values are determined by weighting or averaging the preliminary delta-x and delta-y values. However, the present invention is not so limited.

15 As mentioned above, in another embodiment in accordance with the present invention, an input image is decomposed into frequency band images by recursively applying a Discrete Wavelet Transform. Discrete wavelet transform are known to those of ordinary skill in the  
20 art. This is followed by quantization by thresholding. However, the quantization by thresholding is not a required feature of the present invention. Then, the frequency band images are cross-correlated with reference frequency band images to produce preliminary motion values. The preliminary motion values may be weighted or averaged to produce  
25 final motion values, but the present invention is not so limited.

Figure 4 is a flowchart illustrating a process of detecting motion by decomposing images into frequency bands with a redundant Discrete Wavelet Transform, in accordance with an embodiment of the present invention. A redundant Discrete Wavelet Transform is a Discrete Wavelet Transform performed without down-sampling. In step 510, an input image is received. The image may have been captured by a CMOS sensor, but the present invention is not so limited.

Step 515 is the beginning of a loop for performing a Discrete Wavelet Transform a pre-determined number of times. The loop is exited after the Discrete Wavelet Transform has been performed the pre-determined number of times. The Discrete Wavelet Transform can be performed with many different sets of filter coefficients depending on the type of "wavelet" used. The type of wavelet used determines how the frequency bands are extracted from the image. Embodiments of the present invention are well-suited to all sets of filter coefficients. Moreover, embodiments of the present invention are well-suited to all methods or algorithms of implementing the Discrete Wavelet Transform.

20

In step 520, a Discrete Wavelet Transform is performed on the input image or an image that was previously processed by a Discrete Wavelet Transform, depending on which iteration of the loop is being performed. Each additional Discrete Wavelet Transform decomposes different frequency information from the input image. However, each time the Discrete Wavelet Transform is performed, the same filter coefficients may be used. This greatly simplifies the process versus

25

techniques that use different sets of filter coefficients to extract different frequency information. For example, pipelining is greatly facilitated and circuitry for executing the process is simpler.

5           In step 525, the results of the Discrete Wavelet Transform from step 520 are stored for use as a future reference image. For example, the reference image is used when processing the next input image received in step 510.

10           In step 530, a cross-correlation is performed between the current Discrete Wavelet Transform output and a previously stored reference image from a previous input image.

          In step 535, the result of the cross-correlation is analyzed. Step 540  
15 is a test to determine if the result of the current cross-correlation indicates motion. If it does not, then the process 500 returns to step 515 to perform another Discrete Wavelet Transform. If motion is indicated by the results of the cross-correlation, then the process 500 continues on to step 545.

20

          Step 545 is a test to determine if the result of step 535 indicates a significant deviation from the result of step 535 for the previous iteration of the loop. For example, the cross-correlation result of the third Discrete Wavelet Transform with its reference image is compared to the cross-  
25 correlation result of the second Discrete Wavelet Transform with its reference image. If the current cross-correlation result does not deviate significantly from the previous cross-correlation result, then the process



500 returns to step 515. If there is not a significant deviation, there is not a need to update the motion values. The term "significant deviation" in this context is meant to be flexible in that what is deemed significant may be determined for the particular application in which process 500 is practiced. If there is a significant deviation in the cross-correlation, then the process 500 goes on to step 550 to update the motion values.

Determining if there is a significant deviation in the results allows repetitive features to be removed. For example, if the third cross-correlation result is significantly different from the second cross-correlation result, it may indicate that the second cross-correlation result was affected by repetitive information. The second cross-correlation may have been affected because of the frequency content of its input images, given the nature of the surface for which the images were captured, and the velocity and direction of motion. Therefore, the results from the second cross-correlation are disregarded. If the result of the third Discrete Wavelet Transform is significantly different from the result of the second Discrete Wavelet Transform, then the results of the third Discrete Wavelet Transform are stored as the correct motion values in step 550. In one embodiment, a register containing motion values is updated by overwriting previous motion values.

After updating the motion values in step 550, the process returns to step 515. If the Discrete Wavelet Transform is to be performed again, the process continues at step 520. Otherwise, the process 500 goes to step 555 in which the motion values for this input image are output. Then, the process 500 goes to step 505, in which a new input image is received.

The process 500 then analyzes this input image, with the reference images stored from analyzing the previous input image serving as the reference images.

5           Figure 5A illustrates an exemplary sequence of Discrete Wavelet Transforms and the frequency band images produced therefrom. It will be understood that the example of Figure 5A is exemplary and that many other sequence of Discrete Wavelet Transforms are within the scope of the present invention. Each Discrete Wavelet Transform  
10       produces two or more frequency band images. In this embodiment, each Discrete Wavelet Transform produces two frequency band images. In this embodiment of the present invention, one of the two frequency band images is further processed by an additional Discrete Wavelet Transform. Each successive application of the Discrete Wavelet  
15       Transform extracts a different band of frequency information.

          The first application of the Discrete Wavelet Transform 620a to the input image 605 produces a low-pass and a high-pass frequency band image 625a and 625b. In this embodiment, the low pass frequency band  
20       image 625a is further processed by the second Discrete Wavelet Transform 620b, which produces another low-pass frequency band image 625c and a band-pass frequency band image 625d. The low-pass frequency band image 625c and the band-pass frequency band image 625d comprise frequency bands within the low-pass pass frequency band  
25       image 625a. In this example, the band-pass frequency band image 625d is further processed by the third Discrete Wavelet Transform 620c to produce further band-pass frequency band images 625e and 625f. This

process can continue further to extract further frequency information from the input image. Moreover, the frequency band image that is processed by the next Discrete Wavelet Transform can be tailored to extract the frequency information of interest. For example, rather than,  
5 or in addition to, the second Discrete Wavelet Transform 620b inputting the low-pass frequency band image 625a, the high-pass frequency band image 625b may be processed.

It is not unlikely that the surface being scanned has different  
10 features in the “x” and “y” directions. For example, a wood grained desk may have grain running in the “x” direction. However, many other possibilities exist to produce variations in the features based on orientation. By producing images in which the “x” and “y” components may comprise different frequency bands, embodiments of the present  
15 invention are able to selectively filter information that repeats at a first interval along one of the axes and repetitive information that repeats at a different interval along the other axis. Figure 5B illustrates an exemplary sequence of Discrete Wavelet Transforms and “x” and “y” component frequency band images produced therefrom. It will be  
20 understood that the example of Figure 5B is exemplary and that many other sequence of Discrete Wavelet Transforms are within the scope of the present invention. Each Discrete Wavelet Transform produces four frequency band images, in this embodiment. In this embodiment of the present invention, the four frequency band images is further processed  
25 by four additional Discrete Wavelet Transforms. Each successive application of the Discrete Wavelet Transform extracts a different

frequency band of “x” and “y” component frequency band information, in this embodiment.

The first application of the Discrete Wavelet Transform 620e to the input image 605 produces image “a” 610, image “b” 615, image “c” 620, and image “d” 625. Image “a” 610, has an “x” component low-pass frequency band and a “y” component low-pass frequency band. Image “b” 615, has an “x” component low-pass frequency band and a “y” component high-pass frequency band. Image “c” 620, has an “x” component high-pass frequency band and a “y” component low-pass frequency band. Image “d” 625, has an “x” component high-pass frequency band and a “y” component high-pass frequency band. Thus, the four images (610, 615, 620, 625) produced by the first Discrete Wavelet Transform 620e have “x” and “y” components with different frequency bands.

15

In this embodiment, each of the four images 610, 615, 620, and 625 are processed by an additional Discrete Wavelet Transform 620f-620i, respectively. Each of these additional Discrete Wavelet Transforms produces four images. Thus there are a total of sixteen images 630-645 at this level. It will be understood that the terms “low-pass” frequency band and “high-pass” frequency band in elements 630-645 are relative to the input images 610-625. Thus, Discrete Wavelet Transform 620f inputs image “a” 610 and produces image “a\_1” 630 having “x” and “y” component low-pass frequency bands. These low-pass bands are relative to the low-pass frequency band of image “a” 610. As a further illustration, image “a\_2” 631 has a “y” component high-pass frequency band. It will be understood that this high-pass frequency band is relative

25

to the low-pass frequency band of image “a” 610. Thus, the process illustrated in Figure 5B extracts information for which the “x” and “y” component in a given image may have different frequency bands from each other. Thus, the “x” and “y” directions can be processed in such a fashion to handle cases in which features do not repeat at the same interval along the “x” axis as they do along the “y” axis. The process in Figure 5B can continue further to extract further frequency information from the input image.

10           It is not required that all the images produced by the DWTs are used. For instance, images 610, 634 and 639 could be selected for further processing, while other images are discarded. Therefore, a subset of images may be selected to work with from the images 610-625 from the initial DWT 620e together with or independent of a subset of images 630-15   645 from the next level of DWT processing 620f-i. To save on computation, if any image subset of 610-625 is selected their corresponding DWTs do not need to be performed (e.g., the group 620f-620i DWT). For example, in selecting images 610, 634 and 639, only DWT 620e, 620g and 620h need be performed. DWT 620f is not required since the frequency information 20   decomposed by it is already present in image 610. DWT 620i is not required because the decomposed image bands resulting from it are not used in this example.

Embodiments of the present invention may be stored as instructions on a computer readable medium and executed on a general-purpose processor coupled thereto by a bus. However, the present invention is not limited to software implementation. Embodiments of the

present invention may be implemented as an application specific integrated circuit. Some embodiments of the present invention are implemented by a combination of hardware and software.

- 5           While the present invention has been described in particular embodiments, it should be appreciated that the present invention should not be construed as limited by such embodiments, but rather construed according to the below claims.